



Analyse combinatoire de différents modèles d'urne

En 1923, le mathématicien George Pólya introduit une expérience d'urne aléatoire pour modéliser la propagation d'épidémies. Ce modèle, à base de tirages de boules colorées dans une urne, et ses généralisations ont donné naissance à un grand nombre d'études qui ont conduit à des applications variées, notamment en économie et en finance.

Au milieu des années 2000, le chercheur français Philippe Flajolet propose une nouvelle approche de ces modèles, à base de combinatoire et de séries génératrices. Sa méthode s'applique à la totalité des modèles d'urne dite « équilibrée », là où les techniques de résolution antérieures étaient spécifiques de chaque protocole de tirage.

Cette épreuve est organisée en cinq parties dans une large mesure indépendantes :

- dans la partie I, il est demandé de démontrer un résultat du cours qui est utilisé dans la partie IV ; le résultat de la question 6 sert dans la partie V ;
- la partie II traite un cas particulier qui est généralisé dans la partie IV ;
- la partie III introduit la méthode de Philippe Flajolet ;
- les parties IV et V étudient deux protocoles différents de tirage ; la partie V permet également d'établir un lien avec certaines permutations d'un ensemble fini.

Notations

Dans tout le problème, on définit la famille de polynômes $(L_k)_{k \in \mathbb{N}}$ par

$$\begin{cases} L_0 = 1 \\ L_k = X(X+1)\cdots(X+k-1) \quad \forall k \in \mathbb{N}^* \end{cases}$$

On appelle fonction polynomiale de deux variables réelles u et v , toute combinaison linéaire d'applications de la forme $(u, v) \rightarrow u^i v^j$ où $(i, j) \in \mathbb{N}^2$.

I Résultats préliminaires

I.A – Soit α un réel. On note $f_\alpha : x \mapsto (1-x)^{-\alpha}$.

Q 1. Préciser le domaine de définition D de f_α . Justifier que f_α est de classe C^1 sur D et donner une équation différentielle linéaire du premier ordre vérifiée par f_α sur D .

Q 2. Énoncer le théorème de Cauchy pour une équation différentielle scalaire linéaire du premier ordre et démontrer que, pour tout $x \in]-1, 1[$,

$$f_\alpha(x) = \sum_{n=0}^{+\infty} L_n(\alpha) \frac{x^n}{n!}.$$

Q 3. Rappeler la définition du produit de Cauchy de deux séries entières et énoncer le théorème qui s'y rapporte.

Q 4. En déduire que, pour tout entier n et tous réels α et β ,

$$L_n(\alpha + \beta) = \sum_{k=0}^n \binom{n}{k} L_k(\alpha) L_{n-k}(\beta).$$

I.B –

Q 5. Pour $x \in]-1, 1[$, donner la valeur de la somme de la série entière $\sum_{p=1}^{+\infty} x^p$ ainsi que celle de sa dérivée.

Q 6. Démontrer par récurrence que, pour tout entier $n \in \mathbb{N}^*$, il existe un unique polynôme $R_n \in \mathbb{R}_n[X]$ tel que, pour tout $x \in]-1, 1[$,

$$\sum_{p=1}^{+\infty} p^n x^p = \frac{R_n(x)}{(1-x)^{n+1}}.$$

II Un modèle particulier d'urne de Pólya

On dispose d'un stock infini de boules noires et blanches. Une urne contient initialement une boule noire et une boule blanche. On effectue une suite de tirages selon le protocole suivant :

- on tire au hasard une boule de l'urne ;
- on replace dans l'urne la boule tirée ;
- on ajoute dans l'urne une boule de la même couleur que la boule tirée.

On définit la suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires par $X_0 = 1$ et, pour tout entier $n \geq 1$, X_n donne le nombre de boules blanches dans l'urne après n tirages. On note g_n la fonction génératrice de la variable X_n . On rappelle

que $g_n(t) = \sum_{k=0}^{+\infty} P(X_n = k) t^k$.

Q 7. Déterminer les lois de X_1 , X_2 et X_3 puis les fonctions g_1 , g_2 et g_3 .

Q 8. Soient n et k deux entiers supérieurs ou égaux à 1. Établir que

$$P(X_n = k) = \frac{k-1}{n+1} P(X_{n-1} = k-1) + \frac{n+1-k}{n+1} P(X_{n-1} = k).$$

Q 9. En déduire que, pour tout entier n supérieur ou égal à 1 et tout réel t ,

$$g_n(t) = \frac{t^2 - t}{n+1} g'_{n-1}(t) + g_{n-1}(t).$$

Q 10. Démontrer que, pour tout entier $n \in \mathbb{N}^*$ et tout réel t ,

$$g_n(t) = \frac{1}{n+1} \sum_{k=1}^{n+1} t^k.$$

Q 11. Identifier la loi de X_n et donner son espérance.

III Modèle général d'urne équilibrée

Dans cette partie, on généralise le modèle de la partie précédente.

Soient a_0, b_0, a, b, c et d six entiers naturels. On dispose à nouveau d'un stock infini de boules noires et blanches, mais celles-ci sont cette fois numérotées, à partir de zéro, de manière à pouvoir les différencier. L'urne contient initialement a_0 boules blanches et b_0 boules noires. On effectue une suite de tirages selon le protocole suivant :

- on tire au hasard une boule dans l'urne ;
- on replace cette boule dans l'urne ;
- si la boule tirée est blanche, on ajoute dans l'urne a boules blanches et b boules noires ;
- si la boule tirée est noire, on ajoute dans l'urne c boules blanches et d boules noires.

On suppose que $a + b = c + d$, on dit alors que l'urne est équilibrée et on note $s = a + b = c + d$.

Pour $n \geq 1$, une issue résultant de n tirages successifs est modélisée par le n -uplet indiquant la couleur et le numéro des boules successivement obtenues. On note Ω_n l'ensemble des issues possibles de ces n tirages.

La figure 1 donne deux exemples de 3 tirages ($n = 3$), pour $a_0 = b_0 = 1$, $a = d = 1$ et $b = c = 0$ (modèle de la partie précédente). La boule au dessus de chaque flèche représente celle qui a été tirée.

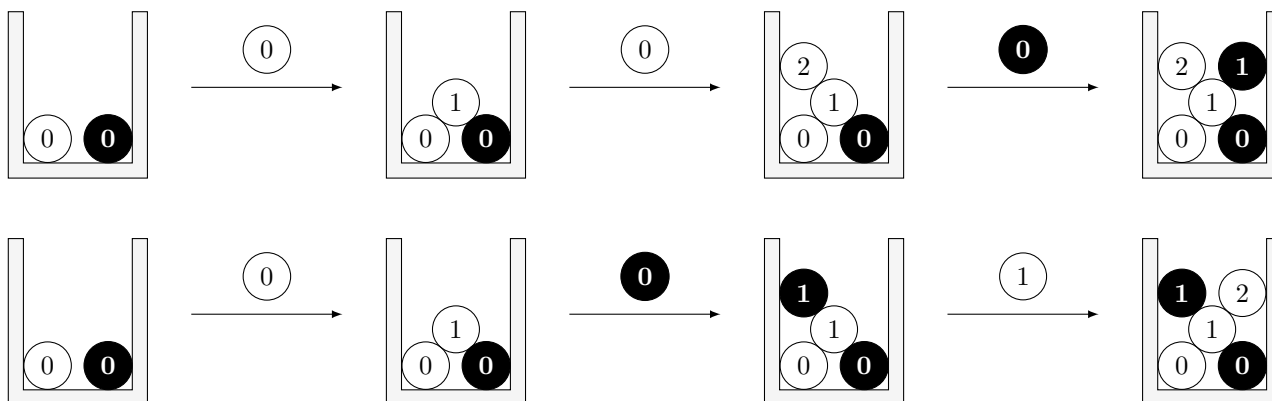


Figure 1 Deux exemples de 3 tirages

La première suite de trois tirages est modélisée par l'issue $\omega_1 = (B_0, B_0, N_0) \in \Omega_3$; la deuxième suite est modélisée par l'issue $\omega_2 = (B_0, N_0, B_1) \in \Omega_3$. On note que ces deux issues différentes aboutissent à la même composition de l'urne.

Pour $\omega \in \Omega_n$ on note $b(\omega)$ (respectivement $n(\omega)$), le nombre de boules blanches (respectivement noires) présentes dans l'urne à la fin des n tirages modélisés par ω .

Pour tous réels u et v , on pose $P_0(u, v) = u^{a_0} v^{b_0}$ et $P_n(u, v) = \sum_{\omega \in \Omega_n} u^{b(\omega)} v^{n(\omega)}$.

Pour $n \geq 1$, on note à nouveau X_n le nombre de boules blanches présentes dans l'urne après n tirages et g_n sa fonction génératrice.

Dans les deux questions suivantes, on suppose $a_0 = 1$, $b_0 = 0$, $a = d = 0$ et $b = c = 1$. En d'autres termes, l'urne contient, au départ, une boule blanche et, à chaque tirage, on ajoute une boule de la couleur opposée à celle qui a été tirée.

Q 12. En dressant la liste de toutes les issues possibles, donner la loi de X_3 .

Q 13. Vérifier que $P_3(u, v) = uv^3 + 4u^2v^2 + u^3v$.

On revient désormais au cas général d'une urne équilibrée.

Q 14. En examinant le nombre de boules dans l'urne juste avant chaque tirage, justifier que, pour $n \geq 1$,

$$\text{card}(\Omega_n) = (a_0 + b_0) \times \cdots \times ((a_0 + b_0 + s(n-1))) = s^n L_n \left(\frac{a_0 + b_0}{s} \right).$$

Q 15. Montrer que, pour tout $n \in \mathbb{N}^*$ et tout $k \in \mathbb{N}$,

$$P(X_n = k) = \frac{\text{card}(\{\omega \in \Omega_n; b(\omega) = k\})}{\text{card} \Omega_n}.$$

Q 16. Justifier les égalités

$$g_n(t) = \frac{1}{\text{card}(\Omega_n)} P_n(t, 1);$$

$$E(X_n) = \frac{1}{\text{card}(\Omega_n)} \frac{\partial P_n}{\partial u}(1, 1).$$

Q 17. Démontrer que, pour tout entier n ,

$$P_{n+1}(u, v) = u^{a+1} v^b \frac{\partial P_n}{\partial u}(u, v) + u^c v^{d+1} \frac{\partial P_n}{\partial v}(u, v). \quad (\text{III.1})$$

Pour tous réels x , u et v , on pose, sous réserve d'existence,

$$H(x, u, v) = \sum_{n=0}^{+\infty} P_n(u, v) \frac{x^n}{n!}.$$

Soit $\rho > 0$. On pose $D_\rho =]-\rho, \rho[\times]0, 2[\times]0, 2[= \{(x, u, v) \in \mathbb{R}^3; |x| < \rho, 0 < u < 2, 0 < v < 2\}$.

Q 18. Justifier que, pour ρ assez petit, la fonction H est bien définie sur D_ρ .

On fixe un tel ρ dans toute la suite de cette partie.

Q 19. Justifier que H admet une dérivée partielle d'ordre 1 par rapport à x sur le domaine D_ρ , obtenue par dérivation terme à terme par rapport à x de l'expression de H .

Q 20. Démontrer que H admet une dérivée partielle d'ordre 1 par rapport à u sur le domaine D_ρ , obtenue par dérivation terme à terme par rapport à u de l'expression de H .

On admet qu'il en est de même pour la variable v .

Q 21. Vérifier que $H(0, u, v) = u^{a_0} v^{b_0}$ puis que H est solution sur D_ρ de l'équation aux dérivées partielles

$$\frac{\partial H}{\partial x}(x, u, v) = u^{a+1} v^b \frac{\partial H}{\partial u}(x, u, v) + u^c v^{d+1} \frac{\partial H}{\partial v}(x, u, v).$$

IV Modèle général d'une urne de Pólya

Dans cette partie, on considère le modèle d'urne équilibrée pour lequel $b = c = 0$. On a donc nécessairement $a = d$. En d'autres termes, à chaque fois qu'on tire une boule, on ajoute a boules de sa couleur dans l'urne. La composition initiale de l'urne est de a_0 boules blanches et b_0 boules noires. Compte tenu des résultats de la partie III, on s'intéresse à l'équation aux dérivées partielles

$$\frac{\partial Y}{\partial x}(x, u, v) = u^{a+1} \frac{\partial Y}{\partial u}(x, u, v) + v^{a+1} \frac{\partial Y}{\partial v}(x, u, v) \quad (\text{IV.1})$$

d'inconnue Y définie sur une partie de $\mathbb{R}^3$ avec la condition $Y(0, u, v) = u^{a_0} v^{b_0}$.

On admet que la fonction G définie sur $U = \{(x, u, v) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^* ; axu^a < 1, axv^a < 1\}$ par

$$G(x, u, v) = u^{a_0} v^{b_0} (1 - axu^a)^{-a_0/a} (1 - axv^a)^{-b_0/a}$$

est solution de l'équation (IV.1).

On reprend la notation de la partie précédente : $D_\rho =]-\rho, \rho[\times]0, 2[$.

Q 22. À l'aide des résultats préliminaires, démontrer qu'il existe $\rho > 0$ tel que $D_\rho \subset U$ et, pour tout $(x, u, v) \in D_\rho$,

$$G(x, u, v) = \sum_{n=0}^{+\infty} Q_n(u, v) \frac{x^n}{n!}$$

où Q_n est une fonction polynomiale de deux variables à préciser.

Q 23. Justifier que G admet une dérivée partielle d'ordre 1 par rapport à x sur le domaine D_ρ , obtenue par dérivation terme à terme par rapport à x de l'expression de H .

Q 24. Démontrer que G admet une dérivée partielle d'ordre 1 par rapport à u sur le domaine D_ρ , obtenue par dérivation terme à terme par rapport à u de l'expression de H .

On admet qu'il en est de même pour la variable v .

Q 25. En déduire que, pour tout entier n , $P_n = Q_n$, puis que H et G coïncident sur D_ρ , où H et $(P_n)_{n \in \mathbb{N}}$ ont été définis dans la partie III.

Q 26. Conclure que, pour tout entier n et pour tout $k \in \llbracket 0, n \rrbracket$,

$$P(X_n = a_0 + ka) = \binom{n}{k} \frac{L_k(a_0/a) L_{n-k}(b_0/a)}{L_n(a_0/a + b_0/a)}.$$

Q 27. À l'aide du résultat précédent, retrouver celui de la question 10.

Q 28. À l'aide des résultats des questions 16 et 19, déterminer l'espérance de X_n .

V Urne de Friedman et montées de permutations

Dans cette partie, on suppose que $a_0 = 1$, $b_0 = 0$, $a = d = 0$, $b = c = 1$, ce qui correspond au protocole utilisé dans les questions 12 et 13 (modèle introduit par Friedman en 1945).

V.A –

On conserve toutes les notations de la partie III (tous les résultats de cette partie peuvent être admis). On a donc en particulier $\text{card}(\Omega_n) = n!$. On admet, pour $0 < u < v$ et $|x|$ assez petit, l'égalité

$$H(x, u, v) = \sum_{n=0}^{+\infty} \frac{x^n}{n!} \left(\sum_{p=1}^{+\infty} p^n (v-u)^{n+1} \left(\frac{u}{v} \right)^p \right).$$

Q 29. À l'aide de la question 6, justifier que, pour tout entier n et tous u et v tels que $0 < u < v$, la somme

$$\sum_{p=1}^{+\infty} p^n (v-u)^{n+1} \left(\frac{u}{v} \right)^p$$

est une fonction polynomiale de u et v .

On a donc, d'après la question 16, pour tout entier n et tout $t \in]0, 1[$, $g_n(t) = \frac{1}{n!} \sum_{p=1}^{+\infty} p^n t^p (1-t)^{n+1}$.

Dans toute la suite, on fixe un entier $n \geq 2$.

Q 30. Montrer que $\sum_{p=n+1}^{+\infty} p^n t^p (1-t)^{n+1} = O(t^{n+1})$.

Q 31. En utilisant ce qui précède et en développant $(1-t)^{n+1}$, déterminer le développement limité de g_n à l'ordre n en 0.

Q 32. En déduire que, pour tout m dans $\llbracket 1, n \rrbracket$,

$$P(X_n = m) = \frac{1}{n!} \sum_{k=0}^{m-1} (-1)^k \binom{n+1}{k} (m-k)^n.$$

V.B – Montées d’une permutation

Soit $n \in \mathbb{N}^*$ et $u = (u_0, \dots, u_n)$ une suite finie d’entiers, deux à deux distincts. Une montée (respectivement descente) de u est un indice $i \in \llbracket 0, n-1 \rrbracket$ tel que $u_i < u_{i+1}$ (respectivement $u_i > u_{i+1}$).

Q 33. Soit k un entier supérieur ou égal à 1, $(u_0, \dots, u_k)$ une suite finie d’entiers, et a un entier tel que $a > u_p$ pour tout $p \in \llbracket 0, k \rrbracket$. On insère la valeur a dans cette suite juste après u_i , avec $i \in \llbracket 0, k-1 \rrbracket$, de manière à obtenir la suite $(u_0, \dots, u_i, a, u_{i+1}, \dots, u_k)$. Comparer le nombre de montées et de descentes de la nouvelle suite par rapport à l’ancienne. On distinguera deux cas.

On note S_n l’ensemble des permutations de $\llbracket 1, n \rrbracket$, c’est-à-dire des fonctions bijectives de $\llbracket 1, n \rrbracket$ dans lui-même. On représente un élément σ de S_n par la suite finie $(\sigma(1), \sigma(2), \dots, \sigma(n))$ et on appelle montée (respectivement descente) de σ une montée (respectivement descente) de cette suite. Par exemple, avec $n = 4$, la permutation σ représentée par la liste $(\sigma(1), \sigma(2), \sigma(3), \sigma(4)) = (4, 1, 3, 2)$ admet une montée et deux descentes. Pour tout entier m , on note $A_{n,m}$ le nombre d’éléments de S_n avec m montées.

Q 34. Déterminer les éléments de S_3 et calculer parmi eux le nombre de permutations avec m montées pour tout entier m . Comparer les valeurs obtenues avec les coefficients de $P_3(X, 1)$ où P_3 a été exprimé à la question 13.

Q 35. Soit $n \geq 2$. Déterminer $A_{n,0}$, $A_{n,n-1}$ et $A_{n,m}$ pour $m \geq n$.

L’objectif de ces dernières questions est de déterminer $A_{n,m}$ pour tous entiers $n \geq 2$ et $m \leq n-1$ en établissant un lien entre ces valeurs et le modèle d’urne étudié dans cette partie.

On étudie un algorithme permettant de construire une permutation de S_n à partir d’une issue correspondant à n tirages.

- On démarre la construction à la suite du premier tirage : on a nécessairement tiré la boule blanche et l’urne contient maintenant une boule de chaque couleur. On considère la suite $(0, 1, 0)$ qui comporte exactement une montée et une descente.
- Si on tire la boule blanche (respectivement noire) lors du deuxième tirage, on insère la valeur 2 au milieu de la première et unique montée (respectivement descente) de la suite pour obtenir la suite $(0, 2, 1, 0)$ (respectivement $(0, 1, 2, 0)$).
- Plus généralement, pour tout $k \in \llbracket 2, n \rrbracket$, si au k -ième tirage on tire une boule blanche (respectivement noire) numérotée p , on insère la valeur k dans la suite au milieu de la $(p+1)$ -ième montée (respectivement descente).
- À la fin de la construction, on supprime les deux 0 de la liste (qui sont nécessairement restés en début et fin de liste). La liste obtenue contient les entiers de 1 à n et représente un élément σ de S_n . Si ω désigne la suite des tirages, on note $\sigma(\omega)$ la permutation obtenue.

À titre d’exemple, construisons $\sigma((B_0, N_0, B_1))$.

— Tirage 1 : B_0

On démarre avec la suite $(0, 1, 0)$.

— Tirage 2 : N_0

L’entier 2 ($k = 2$) s’insère au milieu de la première ($p = 0$) descente (la boule est noire) pour donner la nouvelle suite $(0, 1, 2, 0)$

— Tirage 3 : B_1

L’entier 3 s’insère au milieu de la deuxième montée pour donner $(0, 1, 3, 2, 0)$.

On obtient ainsi $\sigma((B_0, N_0, B_1)) = (1, 3, 2)$.

Q 36. À l’aide de l’algorithme ci-dessus, construire la permutation de S_5 associée à l’issue $(B_0, B_0, N_1, N_0, B_2)$.

Q 37. Réciproquement, soit σ l’élément de S_7 représenté par la suite $(7, 1, 3, 6, 5, 4, 2)$. Déterminer une issue ω comportant 7 tirages telle que $\sigma_\omega = \sigma$.

Q 38. À l’aide de la question 33, comparer, pour une issue quelconque, le nombre de boules blanches dans la composition finale de l’urne au nombre de montées de la permutation qui lui est associée par l’algorithme ci-dessus.

On admet que l’application $\omega \mapsto \sigma(\omega)$ est bijective de Ω_n dans S_n et qu’elle induit une bijection entre l’évènement $(X_n = m)$ et l’ensemble des permutations de S_n ayant $m-1$ montées.

Q 39. Soit $m \in \llbracket 1, n \rrbracket$. Déterminer, pour tout entier $n \geq 2$ et tout $m \in \llbracket 0, n-1 \rrbracket$ le nombre $A_{n,m}$ de permutations de S_n ayant m montées.

• • • FIN • • •
